

December 23, 2014

# Optimal two-qubit tomography based on local and global measurements: Maximal robustness against errors as described by condition numbers

Adam Miranowicz,<sup>1,2</sup> Karol Bartkiewicz,<sup>2,3</sup> Jan Peřina Jr.,<sup>3</sup> Masato Koashi,<sup>4</sup> Nobuyuki Imoto,<sup>5</sup> and Franco Nori<sup>1,6</sup><sup>1</sup>*CEMS, RIKEN, 351-0198 Wako-shi, Japan*<sup>2</sup>*Faculty of Physics, Adam Mickiewicz University, 61-614 Poznań, Poland*<sup>3</sup>*RCPTM, Joint Laboratory of Optics of Palacký University and Institute of Physics of AS CR, Faculty of Science, Palacký University, 17. listopadu 12, 77-146 Olomouc, Czech Republic*<sup>4</sup>*Photon Science Center, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan*<sup>5</sup>*Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan*<sup>6</sup>*Department of Physics, The University of Michigan, Ann Arbor, MI 48109-1040, USA*

We present an error analysis of various tomographic protocols based on the linear inversion for the reconstruction of an unknown two-qubit state. We solve the problem of finding a tomographic protocol which is the most robust against errors in terms of the lowest value (i.e., equal to 1) of a condition number, as required by the Gastinel-Kahan theorem. In contrast, standard tomographic protocols, including those based on mutually unbiased bases, are nonoptimal for determining all 16 elements of an unknown two-qubit density matrix. Our method is based on the measurements of the 16 generalized Pauli operators, where twelve of them can be locally measured, and the other four require nonlocal Bell measurements. Our method corresponds to selectively measuring, one by one, all of the real and imaginary elements of an unknown two-qubit density matrix. We describe two experimentally feasible setups of this protocol for the optimal reconstruction of two photons in an unknown polarization state using conventional detectors and linear-optical elements. Moreover, we define the operators for the optimal reconstruction of the states of multiqubit or multilevel (qudit) systems.

PACS numbers: 03.65.Wj, 03.67.-a, 42.50.Ex

## I. INTRODUCTION

Quantum state tomography (QST) is a method of determining an unknown quantum state (density matrix) in a series of measurements on multiple copies of the state. QST is an essential tool for the verification and benchmarking of quantum devices used, e.g., for quantum state engineering, quantum communication and quantum information processing. Reviews on QST include Refs. [1–3], and more recent results can be found in, e.g., Refs. [4–13] and references therein. Dozens, if not hundreds, of QST protocols have been proposed using various methods both for finite- and infinite-dimensional quantum systems. Among photonic QST protocols, those directly applicable to polarization qubits have attracted considerable interest (for a review see Ref. [14]). Here we analyze mainly two-qubit tomography protocols and describe only their photonic implementations with polarization qubits.

Applied QST is usually based on linear inversion [1] and maximum-likelihood estimation [15–26]. Other proposals of QST are based on, e.g., least-squares inversion [27, 28] (which is also applied in the standard linear-inversion approach to overdetermined systems), as well as Bayesian mean estimation [2, 29, 30], or linear regression estimation [31].

Given this abundance of QST protocols, a natural question would be which of them are optimal, according to some requirements or criteria. The problem of the op-

timality of QST was studied from different perspectives (see, e.g., [4, 5, 9, 29, 32, 33]) including choosing optimal measurement sets to increase the accuracy and efficiency of estimation [34–41]. Various quantitative approaches, in addition to the above references, to testing the performances of QST protocols were recently described by Bogdanov *et al.* [42–44].

In this paper we address the question of finding a QST method, based on linear inversion, which is the most robust against errors, as described by the condition number  $\kappa(A)$ , later defined in Eq. (17) via the spectral norm (i.e., the two-norm condition number). This QST approach is based on solving a linear-system problem,

$$Ax = b, \quad (1)$$

where  $A$  is called here the *rotation matrix* but is also referred to as the *coefficient matrix* or *data matrix* in more mathematical contexts. Moreover,  $b$  is the *observation vector*, which contains the measured data, and  $x = \text{vec}(\rho)$  is a real vector describing the unknown state  $\rho$  to be reconstructed.

Condition numbers are standard parameters characterizing the error stability of, e.g., numerical algorithms [45–47], which, in particular, can be applied to QST based on linear inversion [41, 42]. The significance of applying a condition number in the error analysis of linear systems explains the Gastinel-Kahan theorem [48], which states that the relative distance of a nonsingular square matrix  $A$  to the set of singular matrices is given by the reciprocal

TABLE I: Comparison of error robustness for various two-qubit QST protocols.

Protocol	Based on	Projectors	Number of projectors	local/global projectors	condition no. $\kappa(C) = \kappa^2(A)$	$\min[\text{svd}(C)]$	Eqs.
1	optimal GPOs	$\gamma_k$	16	local & global	1	1	(20), (21)
2	Pauli operators	$\sigma_k \otimes \sigma_l$	16	local	2	1	(A2)
3	James <i>et al.</i> basis [15]	$ \psi_k^{(3)}\rangle$	16	local	60.1	0.1	(A3)
4	standard separable basis [38, 52]	$ \psi_k^{(4)}\rangle$	36	local	9	1	(A4)
5	mutually unbiased bases [39, 50]	$ \psi_k^{(5)}\rangle$	20	local & global	5	1	(A5)
6	Gell-Mann GPOs	$\Gamma_k^{(6)}$	16	local & global	2	$\frac{1}{2}$	(A10)
7	Patera-Zassenhaus GPOs	$\Gamma_k^{(7)}$	16	local & global	2	4	(A11)

of the condition number. Thus, in particular, the condition number  $\kappa(A)$  is a measure of the QST robustness to errors in the observation vector  $b$ . The smaller is the condition number the more robust is the QST method, and the optimal method is described by  $\kappa(A) = 1$ .

To show the importance of error analysis in solving linear systems  $Ax = b$ , let us analyze the following simple example:

$$A = \begin{bmatrix} 6 & 7 \\ 5 & 6 \end{bmatrix} \Rightarrow A^{-1} = \begin{bmatrix} 6 & -7 \\ -5 & 6 \end{bmatrix}. \quad (2)$$

Then for two slightly different observation vectors  $b$  one finds two distinct solutions:

$$b = \begin{bmatrix} .7 \\ .6 \end{bmatrix} \Rightarrow x = \begin{bmatrix} 0 \\ .1 \end{bmatrix} \text{ \& } b = \begin{bmatrix} .71 \\ .59 \end{bmatrix} \Rightarrow x = \begin{bmatrix} .13 \\ -.01 \end{bmatrix}.$$

It is clearly seen that these small relative changes in the observation vector  $b$  are amplified by one order in the solution  $x$ . This unstable solution is a result of an ill-conditioned linear system, as revealed by large condition numbers (as defined below). For example, the condition number based on the spectral norm is equal to  $\kappa(A) \approx 146$ . In general, the solution of a linear system  $Ax = b$  is most stable against changes (errors) in  $b$  if a condition number [say  $\kappa(A)$ ] is equal to 1.

In this paper we describe an optimal two-qubit QST (referred to as Protocol 1) based on both local and global measurements to determine the mean values of some properly chosen generalized Pauli operators (GPOs). These operators, except four diagonal ones, correspond to the Gell-Mann operators for the special unitary group

SU(4) (see, e.g., Ref. [49]). The optimality of Protocol 1 refers to the optimal value (i.e., equal to 1) of the condition number based on the spectral norm.

In Table I, we compare Protocol 1 with six other QST protocols. In particular, we studied a QST method (referred to here as Protocol 5) based on mutually unbiased bases (MUB) [34, 36, 39, 50, 51], which requires both local and global measurements as Protocol 1. Surprisingly, Protocol 5 is five times more sensitive to errors than Protocol 1, and two-and-a-half times worse than the QST based on the local measurements of tensor products of the standard Pauli operators (referred to here as Protocol 2).

Furthermore, we describe two feasible experimental setups for performing Protocol 1 for the optimal reconstruction of an unknown polarization state of two photons.

This paper is organized as follows: In Sec. II, the QST method based on a linear inversion is recalled. Section III introduces the concept of error analysis based on condition numbers for QST. In Sec. IV, the optimal nonlocal tomography based on the measurements of GPOs is proposed as Protocol 1. In Sec. V, two setups of photonic implementations of Protocol 1 are described. In Sec. VI, we show how to construct the operators for the optimal QST of the states of multiqubit and multilevel (qudit) systems. A comparison of Protocol 1 with some other QST methods is presented in the concluding Sec. VII. In Appendix A, the GPOs and projectors of all the discussed two-qubit protocols are summarized. Beam-splitter transformations of entangled projectors are described in Appendix B.

## II. PRINCIPLES OF QST BASED ON LINEAR INVERSION

The numerical procedure to reconstruct a density matrix  $\rho$  from experimental data has been widely used in quantum state engineering (see, e.g., Ref. [2] and references therein).

First, it is useful to represent  $\rho$  as a vector. This procedure can be seen as representing an operator  $\rho$  in Hilbert space as a superoperator  $x = \text{vec}(\rho)$  in Liouville space. The matrix-to-vector operation, which we denote by  $\text{vec}(\rho)$ , can be given by an arbitrarily chosen reordering of the elements  $\rho_{ij}$  of  $\rho$ . For example, one can choose the standard order of  $\rho_{ij}$ , i.e.,  $x' = \text{vec}'(\rho) =$

$[\rho_{00}, \rho_{01}, \dots, \rho_{32}, \rho_{33}]^T$  for a two-qubit state  $\rho$ . The above vector is complex and contains redundant information as  $\rho_{ij} = \rho_{ji}^*$ . Thus, it is convenient to transform  $\rho$  into a real vector, e.g., as follows:

$$x = \text{vec}(\rho) = [\rho_{00}, \text{Re}\rho_{01}, \text{Im}\rho_{01}, \text{Re}\rho_{02}, \text{Im}\rho_{02}, \dots, \rho_{33}]^T, \quad (3)$$

where only the elements  $\rho_{ij}$  for  $i \leq j$  are included. Obviously, any other ordering can be applied but it should be used consistently.

To find all the elements of  $x \equiv [x_i]_{16 \times 1}$ , one has to solve the set of linear equations, given in Eq. (1), where now the rotation matrix is  $A \equiv [A_{ji}]_{N_{\text{eqs}} \times 16}$  and the observation vector is  $b \equiv [b_j]_{N_{\text{eqs}} \times 1}$ . Specifically, the element  $A_{ji}$  is the coefficient of  $x_i$  in the  $j$ th equation for a chosen tomographic rotation, while a given element  $b_j$  of the observation vector  $b$  can correspond, e.g., to coincidence photocounts in optical experiments or the integrated area of spectra in the spectroscopy of nuclear magnetic resonance (NMR). Let us assume that there are  $N_r$  readouts and each of them yields  $N_{\text{vals}}$  values, which can correspond to, e.g., coincidence counts in photon detectors or the number of peaks in the real and imaginary parts of an NMR spectrum. Then the number of equations,  $N_{\text{eqs}}$ , is equal to  $N_r \times N_{\text{vals}}$ . Formally, extra equations can be added, which correspond to the normalization condition,  $\text{Tr}\rho = 1$ .

This problem is usually *overdetermined* if there are more equations than unknowns. The redundant expressions can (sometimes) enable more accurate reconstruction of  $x$ . By applying standard least-squares-fitting analysis one obtains

$$Cx = \tilde{b} \quad \text{with} \quad C = A^\dagger A, \quad \tilde{b} = A^\dagger b, \quad (4)$$

where the overdeterminacy is removed as  $\tilde{b} \equiv [\tilde{b}_j]_{16 \times 1}$  and  $C \equiv [C_{ij}]_{16 \times 16}$ . The matrix  $C$  is sometimes referred to as the *error matrix* [53]. Thus, to reconstruct a density matrix  $\rho$ , it is enough to calculate

$$x = C^{-1}\tilde{b} \rightarrow \rho = \text{vec}^{-1}(x), \quad (5)$$

where  $\text{vec}^{-1}(x)$  is the operation inverse to  $\text{vec}(x)$ . The least-squares analysis is based on the minimalization of  $\chi^2 = \|Ax - b\|^2$ .

### III. ERROR ANALYSIS OF QST BASED ON LINEAR INVERSION

Here let us address the question of how the experimental errors are magnified through the numerical procedure of linear inversion. Thus, the problem now is about the reliability of the reconstructed density matrix  $\rho$  corresponding to the vector  $x = A^{-1}b$  for a given set of rotations  $A$  (representing our linear tomographic system) and for the measured data  $b$ .

Even a simple application of a singular-value decomposition of a nonsingular square matrix  $A \in \mathcal{R}^{n \times n}$ ,

$$A = UDV^T = \sum_{i=1}^n u_i \bar{\sigma}_i v_i^T, \quad (6)$$

implies that the error robustness can be related to the minimal singular value,  $\min_i(\bar{\sigma}_i) \equiv \sigma_{\min}(A)$ . This can be seen by the expansion of the solution  $x$  [47]:

$$x = A^{-1}b = (VD^{-1}U^T)b = \sum_{i=1}^n \frac{u_i^T b}{\bar{\sigma}_i} v_i. \quad (7)$$

Here,  $U = [u_1, \dots, u_n]$  and  $V = [v_1, \dots, v_n]$  are the left- and right-hand singular vectors for  $A$ , respectively, and  $D = \text{diag}([\bar{\sigma}_1, \dots, \bar{\sigma}_n])$  is a diagonal matrix of the singular values  $\bar{\sigma}_i$  for  $A$  (which should not be confused with  $\sigma_i$  denoting the Pauli operators). Thus, by assuming  $\sigma_{\min} \ll 1$ , small errors in  $A$  or  $b$  can induce relatively large errors in  $x$ .

As an indicator of the error robustness (or error sensitivity) of QST methods we apply the *condition number*, which is defined for a nonsingular square matrix  $A$  as follows [45–47]:

$$\text{cond}_{\alpha,\beta}(A) = \|A\|_{\alpha,\beta} \|A^{-1}\|_{\beta,\alpha}, \quad (8)$$

where the convention is used that  $\text{cond}_{\alpha,\beta}(A) = +\infty$  for a singular matrix  $A$ . The subordinate matrix norm  $\|\cdot\|_{\alpha,\beta}$  in Eq. (8) can be given by the vector norms:

$$\|A\|_{\alpha,\beta} = \max_{x \neq 0} \frac{\|Ax\|_\beta}{\|x\|_\alpha}. \quad (9)$$

The condition number was introduced by Turing [54] for the Frobenius norm, but it clearly depends on the underlying norm. Note that the property

$$\text{cond}_{\alpha,\beta}(A) \geq 1 \quad (10)$$

holds for any norm. The condition number has an algebraic interpretation as a normalized Fréchet derivative of the map  $A \rightarrow A^{-1}$  [47].

The importance of  $\text{cond}_{\alpha,\beta}$  in the linear-system error-robustness analysis is based on the Gastinel-Kahan theorem [48] (see also Refs. [45–47]), which states that the relative distance of a nonsingular square matrix  $A$  to the set of singular matrices,

$$\text{dist}_{\alpha,\beta}(A) := \min \left\{ \frac{\|A - P\|_{\alpha,\beta}}{\|A\|_{\alpha,\beta}} : P \text{ is singular} \right\}, \quad (11)$$

is the reciprocal of the condition number,

$$\text{dist}_{\alpha,\beta}(A) = \frac{1}{\text{cond}_{\alpha,\beta}(A)}. \quad (12)$$

In our physical context, the condition number can roughly be interpreted as the rate at which the reconstructed density matrix  $x$  in a given QST method  $Ax = b$

changes with a change in the observation vector  $b$ . If  $\text{cond}_{\alpha,\beta}(A)$  is small, the QST method (or the corresponding rotation matrix  $A$ ) is called *well-conditioned*, which implies that the system is robust against errors in the observation vector  $b$ . However, the problem is referred to as *ill-conditioned* if  $\text{cond}_{\alpha,\beta}(A)$  is large, and *ill-posed* if  $\text{cond}_{\alpha,\beta}(A)$  is infinite. In contrast to a well-conditioned method, an ill-conditioned QST method has a solution  $x$  sensitive to errors (or unstable) in  $b$ , so even a small error in  $b$  can cause a large error in  $x$ .

To show the operational (or physical) importance of condition numbers more explicitly, let us recall a well known theorem (Theorem 8.4 in Ref. [45]): Consider the system given in Eq. (1) with nonsingular  $A$ . Assume perturbations  $\delta b$  in  $b$  and  $\delta A$  in  $A$ , such that  $\|\delta A\| < 1/\|A^{-1}\|$  implying that  $A + \delta A$  is nonsingular. If perturbations  $\delta x$  are defined implicitly by

$$(A + \delta A)(x + \delta x) = b + \delta b, \quad (13)$$

then

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\text{cond}_{\alpha,\beta}(A)}{1 - \text{cond}_{\alpha,\beta}(A) \frac{\|\delta A\|}{\|A\|}} \left( \frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right). \quad (14)$$

By ignoring perturbations in  $A$ , the lower and upper bounds for the relative perturbations in  $x$  are simply given by [45]:

$$\frac{1}{\text{cond}_{\alpha,\beta}(A)} \frac{\|\delta b\|}{\|b\|} \leq \frac{\|\delta x\|}{\|x\|} \leq \text{cond}_{\alpha,\beta}(A) \frac{\|\delta b\|}{\|b\|}, \quad (15)$$

where the right-hand inequality is a special case of the inequality in Eq. (14). Thus, if a condition number  $\text{cond}_{\alpha,\beta}(A)$  is equal (or very close) to one, then small relative changes in the observation vector  $b$  imply equally small relative changes in the reconstructed state  $x$ .

Below, we apply the spectral norm (also called the two-norm) given by the largest singular value of  $A$ , i.e.,

$$\|A\|_{2,2} \equiv \|A\|_2 = \max[\text{svd}(A)] \equiv \sigma_{\max}(A), \quad (16)$$

where the function  $\text{svd}(A)$  returns the singular values of  $A$ . Then the condition number  $\text{cond}_{2,2}(A) \equiv \text{cond}_2(A)$  can be given by a simple formula

$$\kappa(A) \equiv \text{cond}_2(A) = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)}, \quad (17)$$

which is a special case of Eq. (8) because

$$\|A^{-1}\|_2 = \max[\text{svd}(A^{-1})] = \frac{1}{\min[\text{svd}(A)]} \equiv \frac{1}{\sigma_{\min}(A)}. \quad (18)$$

Note that Eq. (17) can be applied not only to square matrices but also to nonsquare ones; e.g., to the rotation matrices  $A$  of the dimensions  $36 \times 16$  and  $20 \times 16$  for Protocols 4 and 5, respectively, as listed in Table I.

Singular values reveal some important aspects of the geometry of a linear transformation  $A$ . In particular,

$\kappa(A)$ , given in Eq. (17) for a square matrix  $A$ , has a clear geometrical interpretation as a degree of the distortion of a unit sphere (or rather hyper-sphere) under the transformation by  $A$  [55]; or, equivalently, as a measure of the elongation of the hyper-ellipsoid  $\{Ax : \|x\|_2 = 1\}$  [47].

One could also calculate the condition number defined via other norms, e.g.,  $\text{cond}_F(A) = \|A\|_F \|A^{-1}\|_F$ , based on the Frobenius norm  $\|A\|_F^2 = \sum_i \sigma_i^2$ . However, for brevity, we apply in this paper only the condition number  $\kappa$ , defined in Eq. (17).

As explained above, the smallest eigenvalue of  $C$  (or  $A$ ) can also be considered an “error robustness parameter”. One can write this parameter as the smallest singular value of  $C$ :

$$\sigma_{\min}(C) = \min[\text{svd}(C)] = \|C^{-1}\|_2. \quad (19)$$

The condition numbers, in contrast to Eq. (19), also contain information about the range of the eigenvalues of  $C$ . In the context of tomographic reconstructions, the parameter  $\sigma_{\min}(C)$  was applied in, e.g., Refs. [53, 56].

One can raise the question of whether a condition number  $\text{cond}_{\alpha,\beta}(A)$  or the minimum singular value  $\sigma_{\min}(A)$  is more appropriate in the analysis of errors in linear systems. Some justifications, like those in Eq. (7) and below Eq. (12), are applicable to both  $\text{cond}_{\alpha,\beta}(A)$  and  $\sigma_{\min}(A)$ . However, inequalities in Eqs. (14) and (15) clearly show the advantage of using  $\text{cond}_{\alpha,\beta}(A)$  over  $\sigma_{\min}(A)$ . Yet another simple argument in support of  $\text{cond}_{\alpha,\beta}(A)$  can be given as follows: Let us rescale vector  $b$  to be ten times its original values. Then  $A$  is also enlarged by 10. This changes  $\sigma_{\min}(A)$ , but the condition number  $\text{cond}_{\alpha,\beta}(A)$  remains unchanged.

#### IV. OPTIMAL NONLOCAL TOMOGRAPHY

Now, let us describe the main result of this paper, i.e., a proposal of an optimal two-qubit QST (referred to as Protocol 1), which is maximally robust against errors, as described by the condition number  $\kappa(A)$  equal to 1.

This protocol requires both local and nonlocal measurements corresponding to measuring the following generalized Pauli operators (GPOs). There are twelve separable (local) GPOs:

$$\begin{aligned} \gamma_1 &= |00\rangle\langle 00|, & \gamma_2 &= |01\rangle\langle 01|, \\ \gamma_3 &= |10\rangle\langle 10|, & \gamma_4 &= |11\rangle\langle 11|, \\ \gamma_5 &= \frac{1}{2}|0\rangle\langle 0| \otimes \sigma_1, & \gamma_6 &= \frac{1}{2}|0\rangle\langle 0| \otimes \sigma_2, \\ \gamma_7 &= \frac{1}{2}\sigma_1 \otimes |0\rangle\langle 0|, & \gamma_8 &= \frac{1}{2}\sigma_2 \otimes |0\rangle\langle 0|, \\ \gamma_9 &= \frac{1}{2}|1\rangle\langle 1| \otimes \sigma_1, & \gamma_{10} &= \frac{1}{2}|1\rangle\langle 1| \otimes \sigma_2, \\ \gamma_{11} &= \frac{1}{2}\sigma_1 \otimes |1\rangle\langle 1|, & \gamma_{12} &= \frac{1}{2}\sigma_2 \otimes |1\rangle\langle 1|, \end{aligned} \quad (20)$$

TABLE II: How to project a given state  $\rho$  of two polarization qubits onto all the *separable* eigenstates  $|\psi_{kl}\rangle$  of the optimal GPOs  $\gamma_k$  ( $k = 1, \dots, 12$ ) in the implementation of Protocol 1 shown in Figs. 1 (Setup 1) and 2 (Setup 2): Rotate locally  $\rho$  by the angles specified below for the HWPs ( $H_1$  and  $H_2$ ) and the QWPs ( $Q_1$  and  $Q_2$ ); and then project the rotated state onto  $|00\rangle \equiv |HH\rangle$ . This probabilistic projection occurs when both detectors  $D_{1H}$  and  $D_{2H}$  ( $D_1$  and  $D_2$ ) click in Setup 1 (Setup 2).

local optimal GPOs	eigenstates $ \psi_{kl}\rangle$ of optimal GPOs	qubit 1		qubit 2	
		$H_1$	$Q_1$	$H_2$	$Q_2$
$\gamma_1$	$ 00\rangle$	0	0	0	0
$\gamma_2$	$ 01\rangle$	0	0	$45^\circ$	0
$\gamma_3$	$ 10\rangle$	$45^\circ$	0	0	0
$\gamma_4$	$ 11\rangle$	$45^\circ$	0	$45^\circ$	0
$\gamma_5$	$ 0+\rangle$	0	0	$22.5^\circ$	0
	$ 0-\rangle$	0	0	$67.5^\circ$	0
$\gamma_6$	$ 0R\rangle$	0	0	0	$45^\circ$
	$ 0L\rangle$	0	0	0	$-45^\circ$
$\gamma_7 = \gamma'_{15}$	$ +\rangle$	$22.5^\circ$	0	0	0
	$ -\rangle$	$67.5^\circ$	0	0	0
$\gamma_8 = \gamma'_{16}$	$ R\rangle$	0	$45^\circ$	0	0
	$ L\rangle$	0	$-45^\circ$	0	0
$\gamma_9$	$ 1+\rangle$	$45^\circ$	0	$22.5^\circ$	0
	$ 1-\rangle$	$45^\circ$	0	$67.5^\circ$	0
$\gamma_{10}$	$ 1R\rangle$	$45^\circ$	0	0	$45^\circ$
	$ 1L\rangle$	$45^\circ$	0	0	$-45^\circ$
$\gamma_{11} = \gamma'_{13}$	$ +\rangle$	$22.5^\circ$	0	$45^\circ$	0
	$ -\rangle$	$67.5^\circ$	0	$45^\circ$	0
$\gamma_{12} = \gamma'_{14}$	$ R\rangle$	0	$45^\circ$	$45^\circ$	0
	$ L\rangle$	0	$-45^\circ$	$45^\circ$	0

TABLE III: How to project  $\rho$  onto all the *entangled* eigenstates  $|\psi_{kl}\rangle$  of the optimal GPOs  $\gamma_k$  ( $k = 13, \dots, 16$ ) in Setup 1: Rotate locally  $\rho$  by the angles specified below for the HWPs and QWPs, and then project the rotated state onto the singlet state  $|\Psi^-\rangle$ . The desired projection is heralded by single clicks in both detectors  $D_{1H}$  and  $D_{2V}$  or  $D_{1V}$  and  $D_{2H}$ .

nonlocal optimal GPOs	eigenstates $ \psi_{kl}\rangle$ of optimal GPOs	qubit 1		qubit 2	
		$H_1$	$Q_1$	$H_2$	$Q_2$
$\gamma_{13}$	$ \Psi^-\rangle$	0	0	0	0
	$ \Psi^+\rangle$	$45^\circ$	$-45^\circ$	0	$45^\circ$
$\gamma_{14}$	$ \bar{\Psi}^-\rangle$	0	$45^\circ$	$-22.5^\circ$	0
	$ \bar{\Psi}^+\rangle$	0	$45^\circ$	$22.5^\circ$	$90^\circ$
$\gamma_{15}$	$ \Phi^-\rangle$	0	$-45^\circ$	0	$45^\circ$
	$ \Phi^+\rangle$	$45^\circ$	0	0	0
$\gamma_{16}$	$ \bar{\Phi}^-\rangle$	0	$45^\circ$	$-22.5^\circ$	$90^\circ$
	$ \bar{\Phi}^+\rangle$	0	$45^\circ$	$22.5^\circ$	0

and four entangled (global) operators

$$\begin{aligned}
\gamma_{13} &= \frac{1}{2}(|\Psi^+\rangle\langle\Psi^+| - |\Psi^-\rangle\langle\Psi^-|), \\
\gamma_{14} &= \frac{1}{2}(|\bar{\Psi}^+\rangle\langle\bar{\Psi}^+| - |\bar{\Psi}^-\rangle\langle\bar{\Psi}^-|), \\
\gamma_{15} &= \frac{1}{2}(|\Phi^+\rangle\langle\Phi^+| - |\Phi^-\rangle\langle\Phi^-|), \\
\gamma_{16} &= \frac{1}{2}(|\bar{\Phi}^+\rangle\langle\bar{\Phi}^+| - |\bar{\Phi}^-\rangle\langle\bar{\Phi}^-|),
\end{aligned} \tag{21}$$

where  $\sigma_n$  are the standard (single-qubit) Pauli operators,  $|\Phi^\pm\rangle = (|00\rangle \pm |11\rangle)/\sqrt{2}$  and  $|\Psi^\pm\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$  are the Bell states,  $|\bar{\Phi}^\pm\rangle = (S \otimes I)|\Phi^\pm\rangle = (|00\rangle \pm i|11\rangle)/\sqrt{2}$  and  $|\bar{\Psi}^\pm\rangle = (S \otimes I)|\Psi^\pm\rangle = (|01\rangle \pm i|10\rangle)/\sqrt{2}$  are Bell-like states, which are given in terms of the phase gate  $S = |0\rangle\langle 0| + i|1\rangle\langle 1|$ . The set of 16 operators is Hermitian and orthogonal in the Hilbert-Schmidt inner product, just as the set  $\{\sigma_1, \sigma_2, \sigma_3, I\}$ . Thus, we refer to the former set as GPOs, although it does not include the identity operator. For clarity, all these 16 GPOs are given explicitly in the standard Fock basis in Appendix A. In terms of the two-qubit density matrix  $\rho$ , each  $\gamma_j$  monitors either the real or imaginary part of a matrix element  $\rho_{kl}$  of the density matrix written in a common basis.

According to the convention, given in Eq. (3), a two-qubit density matrix  $\rho$  can be represented as a real vector  $x = (x_1, \dots, x_{16})$  with its elements given as follows

$$\rho = \begin{bmatrix} x_1 & x_2 + ix_3 & x_4 + ix_5 & x_6 + ix_7 \\ x_2 - ix_3 & x_8 & x_9 + ix_{10} & x_{11} + ix_{12} \\ x_4 - ix_5 & x_9 - ix_{10} & x_{13} & x_{14} + ix_{15} \\ x_6 - ix_7 & x_{11} - ix_{12} & x_{14} - ix_{15} & x_{16} \end{bmatrix}. \tag{22}$$

Then, the mean values  $b_k = \text{Tr}(\rho\gamma_k)$  are simply related to  $x_l$  as

$$\begin{aligned}
b_1 &= x_1 & b_2 &= x_8, & b_3 &= x_{13}, & b_4 &= x_{16}, \\
b_5 &= x_2, & b_6 &= -x_3, & b_7 &= x_4, & b_8 &= -x_5, \\
b_9 &= x_{14}, & b_{10} &= -x_{15}, & b_{11} &= x_{11}, & b_{12} &= -x_{12}, \\
b_{13} &= x_9, & b_{14} &= -x_{10}, & b_{15} &= x_6, & b_{16} &= -x_7.
\end{aligned} \tag{23}$$

It is seen that our method corresponds to selectively measuring, one by one, all the real and imaginary elements of an unknown two-qubit density matrix. Thus, solving this linear problem  $b = Ax$  is trivial because  $A^{-1} = A^T$ ,

where

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & s & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & s & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & s & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & s & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & s & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & s & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (24)$$

with  $s = -1$ . This implies that all the singular values of  $A$  are equal to 1, so the condition number is minimal,  $\kappa(A) = 1$ . For this reason, Protocol 1 is referred to as optimal.

Note that all the nonlocal operators, given in Eq. (21), are related by local operations. For example, they can be expressed in terms of  $\gamma_{13}$  as follows:

$$\begin{aligned} \gamma_{14} &= (S \otimes I) \gamma_{13} (S^\dagger \otimes I), \\ \gamma_{15} &= (I \otimes \sigma_1) \gamma_{13} (I \otimes \sigma_1), \\ \gamma_{16} &= (S \otimes \sigma_1) \gamma_{13} (S^\dagger \otimes \sigma_1), \end{aligned} \quad (25)$$

where  $S = |0\rangle\langle 0| + i|1\rangle\langle 1|$  is the phase gate, and  $I$  is the single-qubit identity operator. More importantly, they can be disentangled by applying the controlled-NOT (CNOT) gate,  $U_{\text{CNOT}}$  and changed into some local GPOs, given in Eq. (20) as follows

$$U_{\text{CNOT}} \gamma_k U_{\text{CNOT}} = \gamma_{k'}, \quad (26)$$

where  $(k, k') = (13, 11), (14, 12), (15, 7),$  and  $(16, 8)$ . All the 28 eigenstates (projectors) of the optimal GPOs are listed in Tables II and III. In particular, those for  $\gamma_5, \dots, \gamma_{12}$  are given by Eq. (20) after applying the eigenstate expansions of the standard Pauli operators

$$\begin{aligned} \sigma_1 &= |+\rangle\langle +| - |-\rangle\langle -| = 2|+\rangle\langle +| - I = I - 2|-\rangle\langle -|, \\ \sigma_2 &= |L\rangle\langle L| - |R\rangle\langle R| = 2|L\rangle\langle L| - I = I - 2|R\rangle\langle R|, \\ \sigma_3 &= |0\rangle\langle 0| - |1\rangle\langle 1| = 2|0\rangle\langle 0| - I = I - 2|1\rangle\langle 1|, \end{aligned} \quad (27)$$

where  $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ ,  $|R\rangle = (|0\rangle - i|1\rangle)/\sqrt{2}$ , and  $|L\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}$ , which can be interpreted, respectively, as diagonal, antidiagonal, right-circular, and left-circular polarization states for the optical polarization qubits.

We note that the number of 28 projectors can be reduced, e.g., by applying the identity resolutions given in

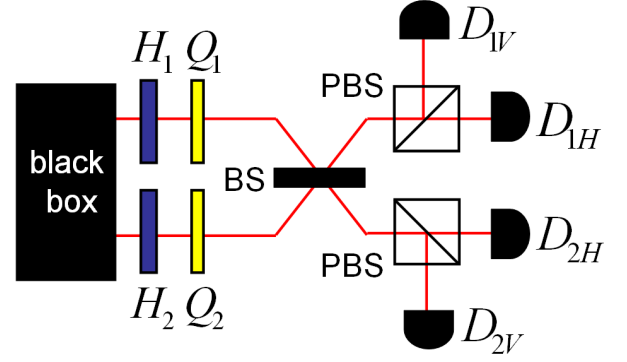


FIG. 1: (Color online) Setup 1 for the experimental implementation of the optimal QST (Protocol 1) of an unknown two-photon polarization state  $\rho$ . Here the state  $\rho$  is the output of a “black box” system, PBSs denote polarizing beam splitters,  $D_{1p}$  and  $D_{2p}$  (with  $p = H, V$ ) correspond to detectors, whose outputs are connected to a coincidence counter (for simplicity, not plotted here). Moreover,  $Q_1$  and  $Q_2$  denote the quarter-wave plates (QWPs) and  $H_1$  and  $H_2$  stand for the half-wave plates (HWPs). A balanced (50 : 50) nonpolarizing beam splitter (BS) is used for the Bell measurement of the nonlocal projectors  $\gamma_n$  for  $n = 13, \dots, 16$ , given in Eq. (21). Here we assume that this BS is removed if the local projectors  $\gamma_n$  for  $n = 1, \dots, 12$ , as listed in Table II, are measured. This method formally corresponds to rotating  $\rho$  by the HWPs and QWPs at the angles specified in Tables II and III; and then projecting them at the fixed states  $|\psi_{\text{fixed}}\rangle = |00\rangle$  (if the BS is removed) and  $|\psi_{\text{fixed}}\rangle = |\Psi^-\rangle$  (if the BS is inserted), respectively. This approach formally corresponds to projecting  $\rho$  onto all the 28 eigenstates  $|\psi_{kl}\rangle$ , from which the mean values of the 16 optimal GPOs,  $\gamma_k$ , can be directly calculated. The projection onto the separable state  $|00\rangle$  is heralded by the coincidence clicks in the detectors  $D_{1H}$  and  $D_{2H}$ , while the projection onto the singlet state  $|\Psi^-\rangle$  occurs for the coincidence clicks in either pair of the detectors:  $D_{1H}$  and  $D_{2V}$  or  $D_{1V}$  and  $D_{2H}$ .

Eq. (27). However, we use this complete set of eigenstates for the same reason of improved experimental stability, as in the case of the application of the standard separable QST (Protocol 4) based on the projections onto all 36 tensor products of the eigenstates of the standard single-qubit Pauli operators [38, 52].

## V. EXPERIMENTAL SETUPS FOR PHOTONIC IMPLEMENTATIONS OF PROTOCOL 1

Here let us describe how to implement the optimal QST to reconstruct an unknown state  $\rho$  of two photons by using polarization degrees of freedom. Thus, in this section, we assume that the qubit states  $|0\rangle$  and  $|1\rangle$  correspond to the horizontal  $|H\rangle$  and vertical  $|V\rangle$  polarizations, respectively.

The projections of a density matrix  $\rho$  onto all the eigenstates  $|\psi_{kl}\rangle$  (with  $k = 1, \dots, 16$  and  $l = 1$  or  $l = 2$ ) of the optimal GPOs can be realized experimentally using

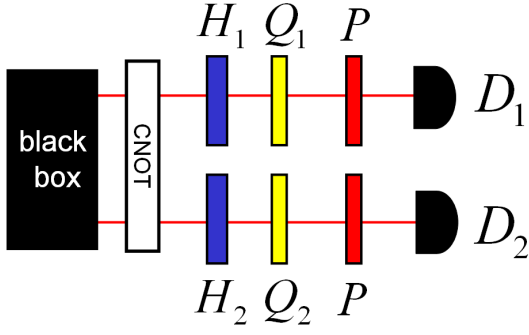


FIG. 2: (Color online) Setup 2 implementing Protocol 1, analogous to Setup 1 in Fig. 1, but here all the rotated states are projected onto the *same* state via the polarizers  $P$  that transmit only photons of one polarization (say, horizontal corresponding to  $|0\rangle \equiv |H\rangle$ ). The CNOT gate is used for disentangling the maximally-entangled projectors  $|\psi_{kl}\rangle$  for  $n = 13, \dots, 16$ , given in Eq. (21). This CNOT should be removed for measuring the local projectors  $\gamma_n$  for  $n = 1, \dots, 12$ , given in Eq. (20). In the latter case, this setup reduces to the standard QST setup of, e.g., Ref. [15]. The efficiency of the setup can be improved if these two polarizers are replaced by PBSs and two extra detectors are placed at the second outputs of the PBSs, as in Setup 1. Anyway, this setup is formally simpler but practically more challenging than Setup 1 because of the use of the optical CNOT gate.

Setup 1, shown in Fig. 1 and described in its caption. The angles of the half-wave plates (HWPs),  $H_1$  and  $H_2$ , as well as the quarter-wave plates (QWPs),  $Q_1$  and  $Q_2$ , are given explicitly in Tables II and III. This setup also includes a removable balanced (50 : 50) beam splitter (BS).

The basic idea is not to directly project  $\rho$  onto  $|\psi_{kl}\rangle$ , but first to rotate  $\rho$  by the HWPs and QWPs, and only then to project them onto some chosen (and fixed) states, e.g., either onto  $|00\rangle \equiv |HH\rangle$  if the BS is removed or the singlet state  $|\Psi^-\rangle$  if the BS is inserted in Setup 1. Then, the mean values of all the optimal GPOs,  $\gamma_k$ , can be calculated directly.

The actions of the HWP and QWP can be defined as

$$H(\theta) = \begin{bmatrix} c & s \\ s & -c \end{bmatrix}, \quad Q(\theta) = \frac{1}{\sqrt{2}} \begin{bmatrix} i+c & s \\ s & i-c \end{bmatrix}, \quad (28)$$

given in terms of  $c = \cos(2\theta)$  and  $s = \sin(2\theta)$ . Note that the operation inverse to the QWP is simply given by  $Q^\dagger(\theta) = -Q(\theta + \pi/2)$ . In special cases,  $H(0)$ ,  $H(\pi/8)$ , and  $H(\pi/4)$  correspond to the phase flip, Hadamard, and bit flip (NOT) gates, respectively, while  $Q(0)$  implements the phase gate  $S$  up to an irrelevant global phase  $\phi = -\pi/4$ . Moreover, the circularly-polarized states can be generated from the horizontally-polarized state as  $|R\rangle = -iQ(\pi/4)|H\rangle$  and  $|L\rangle = -iQ(-\pi/4)|H\rangle$ .

Thus, all the separable eigenstates  $|\psi_{kl}\rangle$  of the optimal GPOs can be transformed into a fixed separable state  $|\psi_{\text{fixed}}\rangle$ , say equal to  $|00\rangle$ , by the local operations implemented by the HWPs and QWPs with the angles speci-

fied in Table II as follows:

$$|\psi_{\text{fixed}}\rangle = U_{kl}|\psi_{kl}\rangle, \quad \text{with } U_{kl} = e^{i\phi_{kl}} Q_1^{kl} H_1^{kl} \otimes Q_2^{kl} H_2^{kl} \quad (29)$$

for  $k = 1, \dots, 12$  and  $l = 1, 2$ . Here we have used the compact notation:  $H_1^{kl} \equiv H_1(\theta = \theta_{H_1}^{kl})$  and  $Q_1^{kl} \equiv Q_1(\theta = \theta_{Q_1}^{kl})$ , etc. Moreover,  $\phi_{kl}$  are irrelevant global phases.

In order to project a given density matrix  $\rho$  onto eight maximally-entangled eigenstates  $|\psi_{kl}\rangle$  of  $\gamma_k$  (for  $k = 13, \dots, 16$ ), one can rotate  $\rho$  by the HWPs and QWPs in eight different ways by the angles specified in, e.g., Table III. Then we can project all of them onto the same maximally-entangled state and to perform its measurement. This Bell measurement can be implemented efficiently using the central BS in Fig. 1 [57].

In Table III, we have assumed that all the eight global projections are locally rotated into the singlet state  $|\Psi^-\rangle$ , which is invariant under the balanced-BS transformation, according to Eq. (B1). The successful singlet-state measurement is heralded by the coincidence counts in either pair of the detectors:  $D_{1H}$  and  $D_{2V}$  or  $D_{1V}$  and  $D_{2H}$ . Alternatively, one can rotate these global projections onto the triplet state  $|\Psi^+\rangle$ . This state after the BS transformation, according to Eq. (B2), is heralded by the coincidence counts in either pair of the detectors:  $D_{1H}$  and  $D_{1V}$  or  $D_{2H}$  and  $D_{2V}$ . In contrast to these two cases, the projections onto the other six entangled states, after the BS transformation, cannot be uniquely distinguished from other orthogonal states in this device, as explicitly given in Eqs. (B3)–(B5).

As explained in the caption of Fig. 1, the desired projections occur probabilistically and they are heralded by proper coincidence counts. Then, the mean values of the GPOs,  $b_k = \text{Tr}(\rho\gamma_k)$ , can directly be obtained from the measured probabilities

$$\text{Tr}(\rho\gamma_k) = \sum_l \lambda_{kl} \langle \psi_{kl} | \rho | \psi_{kl} \rangle = \sum_l \lambda_{kl} \langle \psi_{\text{fixed}} | \rho_{kl} | \psi_{\text{fixed}} \rangle, \quad (30)$$

where  $\rho_{kl} = U_{kl}\rho U_{kl}^\dagger$  is the rotated density matrix  $\rho$ , and  $\lambda_{kl}$  (with  $l = 1$  or  $l = 1, 2$ ) are eigenvalues of  $\gamma_k$ , which can readily be deduced from Eqs. (20), (21), and (27). Thus, the complete Protocol 1 can be applied.

It is worth noting that the standard setup for photonic QST of, e.g., James *et al.* [15] can be used to measure all the 20 local projectors  $|\psi_{kl}\rangle$  (listed in Table II). This would correspond to using Setup 1 without the BS.

Figure 2 shows another setup, which is based on the CNOT gate, or equivalently the controlled-sign (CS) or iSWAP gates. These gates can be used for disentangling all the eight maximally-entangled eigenstates  $|\psi_{kl}\rangle$  (for  $l = 1, 2$ ) into separable eigenstates  $|\psi_{k'l}\rangle$ , i.e.,

$$|\psi_{k'l}\rangle = U_{\text{CNOT}}|\psi_{kl}\rangle \quad (31)$$

for  $k = 13, 14, 15, 16$  and  $k' = 11, 12, 7, 8$ , respectively. It is seen that, contrary to Setup 1, all the rotated states in this setup are projected onto the same separable state  $|00\rangle$ . Thus, this Setup 2 looks formally simpler than

Setup 1. Unfortunately, it is more complicated to be realized practically because of the use of the optical CNOT gate. Ref. [58] lists more than a dozen linear-optical implementations of the *nondestructive* CNOT and/or CS gates including the first proposals of linear-optical entangling gates [59, 60]. The main problem is that these methods usually require some extra resources including ancillae (separable or entangled), feedforward, or extra conventional or even single-photon detectors. Then the complete Setup 2 becomes more complicated than Setup 1. Anyway, Setup 2 shows how the optimal QST can, in principle, be realized also in other systems, where the CNOT gate can be implemented much more efficiently; for example, in nuclear-spin devices using NMR spectroscopy techniques (see, e.g. [1, 61] and references therein). Of course, then the HWPs and QWPs have to be replaced by other feasible local unitary gates, and photocounts will be replaced by the corresponding NMR spectra.

An experimental realization of the proposed optimal QST will be presented elsewhere [62]. Our results will be supported there by a numerical simulation of this experiment assuming realistic single-photon sources including the generation of the vacuum and multiphoton states. Moreover, we can include the effect of imperfect detectors with finite efficiency, dark counts and their limited (i.e., binary) resolution. Such a numerical study can be based on positive-operator-valued measures (POVMs), as applied by us in, e.g., Ref. [63] for a related linear-optical system.

## VI. OPTIMAL TOMOGRAPHY OF QUDIT OR MULTIQUBIT SYSTEMS

Finally, we specify the operators, which should be measured to perform the optimal QST of the density matrix for a  $d$ -level qudit (with  $d = 2, 3, \dots$ ) or  $N$  qubits (where  $d = 2^N$  and  $N = 1, 2, \dots$ ). Again, the optimality of a QST method refers to its maximal robustness against errors, as described by the minimal value of a condition number.

There are  $d^2$  unknown real elements of a state  $\rho$  of a  $d$ -level system, if the normalization of  $\rho$  has to be determined experimentally. Thus, we need the same number of optimal GPOs.

Let us analyze the explicit form, as given in Eq. (A1), of the GPOs for the optimal two-qubit tomography. The symmetry of these two-qubit operators clearly shows how the optimal tomography can be generalized for multi-qubit or multilevel states. Thus, one can deduce that the GPOs (denoted as  $\gamma_n^{[d]}$ ) for the optimal QST of a  $(d \times d)$ -dimensional density matrix can be given as

$$\{\gamma_n^{[d]}; n = 1, \dots, d^2\} = \{X_{k,k}^{[d]}, X_{k<l}^{[d]}, Y_{k<l}^{[d]}; k, l = 0, \dots, d-1\}, \quad (32)$$

where

$$X_{k,l}^{[d]} = \frac{1}{2}(|k\rangle\langle l| + |l\rangle\langle k|), \quad (33)$$

$$Y_{k,l}^{[d]} = \frac{1}{2}(-i|k\rangle\langle l| + i|l\rangle\langle k|). \quad (34)$$

The measurement of the each GPO,  $\gamma_n^{[d]}$ , corresponds to a direct measurement of either the real ( $X_{i,j}^{[d]}$ ) or imaginary ( $Y_{i,j}^{[d]}$ ) part of the element  $\rho_{ij}$  of a given density matrix  $\rho$ . By this construction, it is seen that the QST based on the measurement of  $\gamma_n^{[d]}$  is optimal as described by the condition number  $\kappa(A) = 1$ .

For a comparison, let us analyze another simple approach to QST of  $N$  qubits, which is based on the local measurements of the tensor products of the standard single-qubit Pauli operators, i.e.,

$$\Gamma_n^{[d]} = \sigma_{n_1} \otimes \sigma_{n_2} \otimes \dots \otimes \sigma_{n_N}, \quad (35)$$

where  $n = \{n_1, n_2, \dots, n_N\}$  and  $n_i = 0, 1, 2, 3$ . This multiindex  $n$  can be considered a single number written in the ternary numeral system, i.e.,  $n = 1 + \sum_{i=1}^N 4^{N-i} n_i$ . The condition number for this method is given by  $\kappa(C) = \kappa^2(A) = 2$ . Thus, it is seen that this simple approach, based on local measurements, is less robust against errors in comparison to our approach, based on both local and global measurements. Moreover, the latter method can be applied for the QST of a single  $d$ -level system only if  $d = 2^N$ . Note that our GPO-based method can be used for qudits with any number of levels even if  $d \neq 2^N$ .

### A. Single-qubit tomography

For clarity and completeness of our presentation, let us analyze the simplest case of the reconstruction of a single-qubit density matrix,

$$\rho = \begin{bmatrix} x_1 & x_2 + ix_3 \\ x_2 - ix_3 & x_4 \end{bmatrix}. \quad (36)$$

By applying the optimal projectors  $\gamma_n^{[1]}$ , given by

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad (37)$$

one finds that the protocol is optimal, as described by

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad (38)$$

with the condition number  $\kappa(A) = 1$ .

For a comparison, let us analyze the standard approach of finding four unknowns  $x_n$  (for  $n = 1, \dots, 4$ ) by applying the Pauli operators together with the identity operator,  $\{\sigma_1, \sigma_2, \sigma_3, I\}$ , which is a special case of Eq. (35). A



simple calculation shows that

$$A = \begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad (39)$$

which leads to the condition number  $\kappa(C) = \kappa^2(A) = 2$ . Thus, this approach is not optimal.

Note that this standard approach is optimal for finding only three unknowns  $x_n$  (for  $n = 1, 2, 3$ ), as given by Eq. (36), where  $x_4$  is determined from the normalization condition, as  $x_4 = 1 - x_1$ , and the projectors include only the Pauli operators without the identity operator. In this case, one finds that

$$A = 2 \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}, \quad (40)$$

for which the condition number is  $\kappa(A) = 1$ , as desired. In this case, the corresponding observation vector  $b$  should be displaced as,  $b \rightarrow b + [0, 0, 1]$ .

In the concluding section, we address the problem of reducing the number of variables by applying the normalization condition and its effect on the error robustness of two-qubit tomography.

## VII. DISCUSSION AND CONCLUSIONS

The main problem studied here was to find a QST method which is the most robust against errors, as described by the condition numbers defined via the spectral norm (i.e., the two-norm condition number).

If QST is directly based on solving a linear-system problem, given in Eq. (1), then the condition number  $\kappa(A)$  is a good measure of the QST robustness against errors in the observation vector  $b$ . Indeed, according to the Gastinel-Kahan theorem [48], the condition number  $\kappa(A)$  has a clear geometric meaning as the reciprocal of a relative distance of a nonsingular matrix  $A$  to the set of singular matrices. The smaller is the condition number the more robust is the QST method, and the optimal method is described by  $\kappa(A) = 1$ .

The main advantage of using condition numbers to describe the error robustness of QST methods might be that the condition numbers determine lower and upper bounds, as given by Eqs. (14) and (15), on the errors in a reconstructed state  $\rho$ . This estimation of error robustness depends on the errors in the measured data, although the real sources of the errors are irrelevant. For example, they can be related to imperfect detectors, realistic photon sources, lossy and unbalanced linear-optical elements (including beam splitters and wave plates), etc.

We found such a QST method (referred to as Protocol 1), based on the measurement of the generalized Pauli

operators, defined in Eqs. (20)–(21). Protocol 1 corresponds to measuring one by one all of the real and imaginary elements of an unknown two-qubit density matrix. This approach results in the condition number  $\kappa(A) = 1$ . Thus, Protocol 1 can be considered as the most robust against errors, which can occur in the observation vector  $b$ . Moreover, we described two experimentally feasible setups of photonic implementations of Protocol 1.

In Table I, we compared this error robustness of Protocol 1 with six other QST methods: Protocol 2 is based on the measurements of all the 16 tensor products of the standard Pauli operators,  $\sigma_i \otimes \sigma_j$ . Protocol 3 is the well-known QST method of James *et al.* [15] based on 16 projectors given explicitly in Appendix A. Protocol 4, often referred to as standard-separable QST, is based on the measurements of all the 36 eigenstates of the operators  $\sigma_i \otimes \sigma_j$  used in Protocol 2. Protocol 5 is based on the projections onto MUB according to the original idea of Wootters and Fields [34], later studied in, e.g., Refs. [36, 50, 51], and experimentally applied by Adamson and Steinberg [39]. Protocols 6 and 7 are based on the measurements of the Gell-Mann GPOs for the special unitary group SU(4) (see, e.g., Ref. [49]) and the Patera-Zassenhaus GPOs for the general linear group GL(4, C) [64], respectively. The projectors for all these seven protocols are defined explicitly in Appendix A.

It is worth noting that the projections of  $\rho$  onto all the eigenstates of tensor products of the standard Pauli operators  $\sigma_i \otimes \sigma_j$  in Protocol 4 also enable the application of Protocol 2 based on the measurements of  $\sigma_i \otimes \sigma_j$ . Namely, the set of 16 equations in Protocol 2 can be obtained by proper linear combinations of the 36 equations in Protocol 4, according to the eigenstate expansions of the Pauli operators, given in Eq. (27). Thus, the error robustness of QST can be improved 4.5 times if described by the conditions numbers  $\kappa(C) = \kappa^2(A)$  (see Table I). This approach, based on pure-state projections of an unknown state  $\rho$ , can also be applied to measure the GPOs in Protocols 1, 7, and 8.

Protocols 2–5 are based solely on local rotations and local measurements. However, all the other protocols require both single-qubit and nonlocal two-qubit projections. These nonlocal projections of an unknown state onto a given Bell state can be realized effectively in the standard Bell analyzers [1, 57] as applied in Setups 1 and 2 presented in Figs. 1 and 2. Note that Ref. [39] describes not only a proposal to use MUB for tomography (referred to here as Protocol 5), but also reports an experimental photonic implementation, which includes nonlocal projections of a given state onto Bell-like states, which corresponds to Setup 1 in Fig. 1.

Protocols 1 and 6 are apparently similar. Indeed, all twelve nondiagonal Gell-Mann GPOs are the same as the optimal GPOs, i.e.,  $\Gamma_n^{(6)} = \gamma_n$ , for  $n = 5, \dots, 16$ . However, the other four diagonal Gell-Mann GPOs are different from  $\gamma_n$ . Note that the Gell-Mann GPOs, like the standard Pauli matrices, are Hermitian, traceless, and orthogonal in the Hilbert-Schmidt inner product. While

the optimal GPOs  $\gamma_n$  are Hermitian and orthogonal, but the diagonal ones are not traceless. This difference implies that the error robustness of Protocol 6 is twice worse than that of Protocol 1, and this is the same as in Protocol 2 being solely based on local measurements. Thus, the nonlocal-projections in Protocol 6 do not offer any advantage (in terms of the condition number) over the local ones in Protocol 2.

Surprisingly, the MUB-based Protocol 5 is nonoptimal concerning the error robustness, measured by the condition numbers  $\kappa(C) = \kappa^2(A)$ , which is five times worse than the optimal Protocol 1 and  $\frac{5}{2}$  times worse than Protocol 2 based solely on local measurements.

It is quite counterintuitive that the projections onto MUB of Refs. [36, 39, 50, 51] are the nonoptimal choices of measurements in terms of the lowest condition number. Nevertheless, nonoptimality of MUB was also observed in other contexts, e.g., in the detection of the Einstein-Podolski-Rosen steering, where random measurements are in some cases better than MUB (maximally noncommuting observables) [65]. It should be also noted that the MUB-based Protocol 5 is the most robust against errors among the QST protocols (listed in Table I), which are based solely on pure-state projections.

It should be stressed that we discussed the reconstruction of 16 real elements of an unknown two-qubit density matrix  $\rho$ . One could argue that only 15 elements are unknown, since the 16th element can be calculated from the normalization condition  $\text{Tr}\rho = 1$ . However, in the experiments with imperfect detection efficiency (like typical photon counting), this normalization has to be determined in a separate measurement (corresponding to a separate equation). Thus, one has to determine all the 16 unknown elements in such experiments. In particular, the MUB-based reconstruction of only 15 elements, results in the case of perfect error robustness as described by the condition number  $\kappa(A) = 1$ . However, by including the 16th unknown element of  $\rho$  (say  $\rho_{44}$ ), the error robustness of this MUB approach is five times worse. As explained above, we prefer to reconstruct all the 16 elements for operational reasons. To clarify this point let us give a typical example of a two-photon state, where  $\rho$  is unnormalized. By using the parametrization of  $\rho$ , given in Eq. (22), one can say that  $\rho$  corresponds to  $\eta\rho'$ , where  $\rho'$  is the normalized density operator of the measured two-photon state and  $\eta$  is the unknown efficiency for detecting the two photons. By knowing all the diagonal terms of  $\rho$ , one can directly determine  $\eta$  as  $\text{Tr}\rho$ .

We also generalized our approach by defining observables for the optimal reconstruction of the unknown state of an arbitrary number of qubits or arbitrary-level qudits. This method is the most robust against errors, since  $\kappa(A) = 1$  for any dimension of the state. For a comparison, we analyzed a simple approach to QST of a multiqubit system based on the local measurements of the tensor products of the standard single-qubit Pauli operators. The latter approach is not optimal, as described by the condition number  $\kappa(A) = \sqrt{2}$ .

Finally, we express our hope that the proposed tomographic protocol, which is optimally robust against errors and can be easily implemented by using, e.g., linear optics, can become a useful tool for quantum engineering and quantum information processing.

## Acknowledgments

The authors thank Antonín Černoch, Daoyi Dong, Karel Lemr, Şahin Özdemir, Bo Qi, and Jan Soubusta for discussions. A.M. is supported by the Polish National Science Centre under Grants DEC-2011/03/B/ST2/01903 and DEC-2011/02/A/ST2/00305. K.B. acknowledges the support by the Polish National Science Centre (Grant No. DEC-2013/11/D/ST2/02638) and by the Foundation for Polish Science (START Programme). K.B. and J.P. are supported by the project No. LO1305 of the Ministry of Education, Youth and Sports of the Czech Republic. N.I. is supported by JSPS Grant-in-Aid for Scientific Research(A) 25247068. F.N. is partially supported by the RIKEN iTHES Project, MURI Center for Dynamic Magneto-Optics, and a Grant-in-Aid for Scientific Research (S).

## Appendix A: Projectors for quantum-state-tomography protocols

For the benefit of the reader, we explicitly show here the QST projectors and other details for the QST protocols discussed in Table I.

### Protocol 1 with optimal generalized Pauli operators

The two-qubit optimal GPOs, given by Eqs. (20) and (21), have the following symmetrical forms in the stan-

dard computational basis:

$$\begin{aligned}
\gamma_1 &= \text{diag}([1, 0, 0, 0]), & \gamma_2 &= \text{diag}([0, 1, 0, 0]), \\
\gamma_3 &= \text{diag}([0, 0, 1, 0]), & \gamma_4 &= \text{diag}([0, 0, 0, 1]), \\
\gamma_5 &= \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, & \gamma_6 &= \frac{1}{2} \begin{bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\
\gamma_7 &= \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, & \gamma_8 &= \frac{1}{2} \begin{bmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\
\gamma_9 &= \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, & \gamma_{10} &= \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix}, \\
\gamma_{11} &= \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, & \gamma_{12} &= \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{bmatrix}, \\
\gamma_{13} &= \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, & \gamma_{14} &= \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\
\gamma_{15} &= \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, & \gamma_{16} &= \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}.
\end{aligned} \tag{A1}$$

For this protocol, the condition numbers are minimal,  $\kappa(C) = \kappa(A) = 1$ .

#### Protocol 2 with standard Pauli operators

Protocol 2 for two-qubit QST is based on measuring all the tensor products of the single-qubit Pauli operators (see, e.g., Ref. [15] and references therein):

$$\Gamma_{4i+j+1}^{(2)} = \sigma_i \otimes \sigma_j \tag{A2}$$

for  $i, j = 0, \dots, 3$ , where  $\sigma_0 = I$  is the identity operator. This is a natural generalization of the single-qubit QST. For this protocol, the condition numbers are  $\kappa(C) = \kappa^2(A) = 2$ .

#### Protocol 3 of James et al.

Protocol 3 for QST is based on the following projections  $\{|\psi_n^{(3)}\rangle\}$ , which were applied in the QST experiment

performed by James et al. [15]:

$$\begin{aligned}
\{|\psi_n^{(3)}\rangle\} &= \{|00\rangle, |01\rangle, |0+\rangle, |0L\rangle, \\
&\quad |10\rangle, |11\rangle, |1+\rangle, |1L\rangle, \\
&\quad |R0\rangle, |R1\rangle, |R+\rangle, |RL\rangle, \\
&\quad |+\rangle, |+\rangle, |+\rangle, |+\rangle, |R\rangle\}. \tag{A3}
\end{aligned}$$

The resulting condition numbers are the largest among the studied protocols, as  $\kappa(C) = \kappa^2(A) \approx 60.1$ .

#### Protocol 4 with Pauli operator eigenstates

Protocol 4 is probably the most popular experimental two-qubit QST method and is referred to as standard-separable QST. It is based on the projections onto all of the 36 tensor products of the eigenstates of the standard single-qubit Pauli operators [52] (see also [38]):

$$\begin{aligned}
\{|\psi_n^{(4)}\rangle\} &= \{|00\rangle, |01\rangle, |10\rangle, |11\rangle, |\pm\rangle, |\pm\rangle, \\
&\quad |0\pm\rangle, |\pm 0\rangle, |1\pm\rangle, |\pm 1\rangle, \\
&\quad |0R\rangle, |R0\rangle, |1R\rangle, |R1\rangle, |0L\rangle, |L0\rangle, |1L\rangle, |L1\rangle, \\
&\quad |R\pm\rangle, |\pm R\rangle, |L\pm\rangle, |\pm L\rangle, \\
&\quad |RR\rangle, |RL\rangle, |LR\rangle, |LL\rangle\}. \tag{A4}
\end{aligned}$$

The corresponding condition numbers are  $\kappa(C) = \kappa^2(A) = 9$ .

#### Protocol 5 based on mutually unbiased bases

Protocol 5 is based on the five MUB of Adamson and Steinberg [39]:

$$\{|\psi_n^{(5)}\rangle\} = \{|\psi_n^A\rangle, |\psi_n^B\rangle, |\psi_n^C\rangle, |\psi_n^D\rangle, |\psi_n^E\rangle\}, \tag{A5}$$

where

$$\begin{aligned}
\{|\psi_n^A\rangle\} &= \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}, \\
\{|\psi_n^B\rangle\} &= \{|R\pm\rangle, |L\pm\rangle\}, \\
\{|\psi_n^C\rangle\} &= \{|\pm R\rangle, |\pm L\rangle\}, \\
\{|\psi_n^D\rangle\} &= \left\{ \frac{1}{\sqrt{2}}(|R0\rangle \pm i|L1\rangle), \frac{1}{\sqrt{2}}(|R1\rangle \pm i|L0\rangle) \right\}, \\
\{|\psi_n^E\rangle\} &= \left\{ \frac{1}{\sqrt{2}}(|RR\rangle \pm i|LL\rangle), \frac{1}{\sqrt{2}}(|RL\rangle \pm i|LR\rangle) \right\}. \tag{A6}
\end{aligned}$$

This protocol was applied in the QST experiment of Ref. [39]. The twenty states of the MUB include twelve separable and eight Bell-like states. The latter are simply related to the standard Bell states by local operations as follows:

$$\begin{aligned}
|\psi_{1,2}^D\rangle &= \frac{1}{\sqrt{2}}(|R0\rangle \pm i|L1\rangle) \cong (SHS \otimes \sigma_2)|\Phi^\pm\rangle, \\
|\psi_{3,4}^D\rangle &= \frac{1}{\sqrt{2}}(|R1\rangle \pm i|L0\rangle) \cong (SHS \otimes \sigma_2)|\Psi^\pm\rangle, \\
|\psi_{1,2}^E\rangle &= \frac{1}{\sqrt{2}}(|RR\rangle \pm i|LL\rangle) \cong (SHS \otimes SH)|\Phi^\pm\rangle, \\
|\psi_{3,4}^E\rangle &= \frac{1}{\sqrt{2}}(|RL\rangle \pm i|LR\rangle) \cong (SHS \otimes SH)|\Psi^\pm\rangle, \tag{A7}
\end{aligned}$$

where  $S$  and  $H$  are the phase and Hadamard gates, respectively, and the sign  $\cong$  indicates that the corresponding expressions are equal up to irrelevant global phase factors.

Another MUB for two qubits was studied by Bandyopadhyay *et al.* [50] and others [36, 51]. Here we rewrite this MUB explicitly in terms of the twelve separable and eight Bell-like states, analogously with Eq. (A6). We have

$$\begin{aligned} \{|\psi_n^B\rangle\} &= \{|\pm+\rangle, |\pm-\rangle\}, \\ \{|\psi_n^C\rangle\} &= \{|RR\rangle, |RL\rangle, |LR\rangle, |LL\rangle\}, \\ \{|\psi_n^D\rangle\} &= \{U_1|\Phi^\pm\rangle, U_1|\Psi^\pm\rangle\} \\ &\cong \left\{\frac{1}{\sqrt{2}}(|L0\rangle \pm |R1\rangle), \frac{1}{\sqrt{2}}(|R0\rangle \pm |L1\rangle)\right\}, \\ \{|\psi_n^E\rangle\} &= \{U_2|\Phi^\pm\rangle, U_2|\Psi^\pm\rangle\} \\ &= \left\{\frac{1}{\sqrt{2}}(|0L\rangle \pm |1R\rangle), \frac{1}{\sqrt{2}}(|0R\rangle \pm |1L\rangle)\right\} \end{aligned} \quad (\text{A8})$$

and the basis  $\{|\psi_n^A\rangle\}$  is the same as in Eq. (A6). Moreover,  $U_1 = SH \otimes I$  and  $U_2 = I \otimes SH$ .

It is easy to confirm that both Eqs. (A6) and (A8) represent MUB because it holds

$$|\langle\psi_m^X|\psi_n^Y\rangle| = \frac{1}{2} \quad (\text{A9})$$

for any  $m, n \in \{1, \dots, 4\}$  and  $X \neq Y \in \{A, \dots, E\}$ .

The condition numbers  $\kappa(C) = \kappa^2(A)$  are equal to 5 for both of these MUB, given by Eqs. (A6) and (A8).

### Protocol 6 with Gell-Mann generalized Pauli operators

Protocol 6 is based on the Gell-Mann GPOs for the special unitary group  $SU(4)$ . These GPOs can be given in the standard computational basis (see, e.g., Ref. [49]) as

$$\begin{aligned} \Gamma_1^{(6)} &= \frac{1}{2}I_4, \\ \Gamma_2^{(6)} &= \frac{1}{2}\text{diag}([1, -1, 0, 0]), \\ \Gamma_3^{(6)} &= \frac{1}{2\sqrt{3}}\text{diag}([1, 1, -2, 0]), \\ \Gamma_4^{(6)} &= \frac{1}{2\sqrt{6}}\text{diag}([1, 1, 1, -3]), \\ \Gamma_n^{(6)} &= \gamma_n \quad \text{for } n = 5, \dots, 16, \end{aligned} \quad (\text{A10})$$

where  $\gamma_n$  are our GPOs, defined in Eq. (A2), and  $I_4$  is the four-dimensional identity operator. The corresponding condition numbers are  $\kappa(C) = \kappa^2(A) = 2$ .

### Protocol 7 with Patera-Zassenhaus generalized Pauli operators

Protocol 7 is based on the Patera-Zassenhaus GPOs for the general linear group  $GL(4, \mathbb{C})$  [64]:

$$\begin{aligned} \Gamma_1^{(7)} &= D, & \Gamma_2^{(7)} &= D^2, & \Gamma_3^{(7)} &= D^3, \\ \Gamma_4^{(7)} &= B, & \Gamma_5^{(7)} &= B^2, & \Gamma_6^{(7)} &= B^3, \\ \Gamma_7^{(7)} &= BD, & \Gamma_8^{(7)} &= BD^2, & \Gamma_9^{(7)} &= BD^3, \\ \Gamma_{10}^{(7)} &= B^2D, & \Gamma_{11}^{(7)} &= B^2D^2, & \Gamma_{12}^{(7)} &= B^2D^3, \\ \Gamma_{13}^{(7)} &= B^3D, & \Gamma_{14}^{(7)} &= B^3D^2, & \Gamma_{15}^{(7)} &= B^3D^3, \end{aligned} \quad (\text{A11})$$

and  $\Gamma_{16}^{(7)} = I_4$ , where

$$\begin{aligned} D &= \exp(i\pi/4) \text{diag}([1, i, -1, -i]), \\ B &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{bmatrix}. \end{aligned} \quad (\text{A12})$$

The corresponding condition numbers for this protocol are the same as for Protocols 2 and 7, i.e.,  $\kappa(C) = \kappa^2(A) = 2$ .

### Appendix B: Beam-splitter transformation of entangled projectors

Here we show explicitly the transformation via a 50 : 50 beam splitter of the Bell and Bell-like states, which are the entangled projectors, i.e., the eight maximally-entangled eigenstates of the optimal GPOs  $\gamma_n$  for  $n = 13, \dots, 16$ .

The 50 : 50 beam-splitter transformation  $U_{BS}$  can be implicitly given by the transformation between the input ( $a_{1p}$  and  $a_{2p}$ ) and output ( $b_{1p}$  and  $b_{2p}$ ) annihilation operators, e.g., [57, 66]:  $a_{1p} = (b_{1p} + b_{2p})/\sqrt{2}$  and  $a_{2p} = (b_{1p} - b_{2p})/\sqrt{2}$ , for two polarizations  $p = H, V$ . Then the entangled projectors are transformed as follows

$$U_{BS}|\Psi^-\rangle = -|\Psi^-\rangle, \quad (\text{B1})$$

$$U_{BS}|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|HV, \text{vac}\rangle - |\text{vac}, HV\rangle), \quad (\text{B2})$$

$$\begin{aligned} U_{BS}|\Phi^\pm\rangle &= \frac{1}{2}(|2H, \text{vac}\rangle - |\text{vac}, 2H\rangle) \\ &\quad \pm \frac{1}{2}(|2V, \text{vac}\rangle - |\text{vac}, 2V\rangle), \end{aligned} \quad (\text{B3})$$

$$\begin{aligned} U_{BS}|\bar{\Psi}^\pm\rangle &= c^\pm(|HV, \text{vac}\rangle - |\text{vac}, HV\rangle) \\ &\quad - c^\mp(|H, V\rangle - |V, H\rangle) \\ &= \sqrt{2}c^\pm U_{BS}|\Psi^\pm\rangle - \sqrt{2}c^\mp |\Psi^\mp\rangle, \end{aligned} \quad (\text{B4})$$

$$\begin{aligned} U_{BS}|\bar{\Phi}^\pm\rangle &= \frac{1}{2}(|2H, \text{vac}\rangle - |\text{vac}, 2H\rangle) \\ &\quad \pm \frac{i}{2}(|2V, \text{vac}\rangle - |\text{vac}, 2V\rangle), \end{aligned} \quad (\text{B5})$$

where  $|\text{vac}\rangle$  stands for the vacuum,  $c^\pm = (1 \pm i)/(2\sqrt{2})$ ,  $|V, H\rangle = b_{1V}^\dagger b_{2H}^\dagger |\text{vac}, \text{vac}\rangle$ , and  $|2V, \text{vac}\rangle =$

$\frac{1}{\sqrt{2}} b_{1V}^{\dagger 2} |\text{vac}, \text{vac}\rangle$ , etc.

- 
- [1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2001).
- [2] M. G. A. Paris and J. Řeháček (eds.), *Quantum State Estimation*, Lecture Notes in Physics, Vol. 649 (Springer, Berlin, 2004).
- [3] G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, "Quantum Tomography," *Advances in Imaging and Electron Physics* **128**, 205-308 (2003).
- [4] G. M. D'Ariano and P. Perinotti, "Optimal data processing for quantum measurements," *Phys. Rev. Lett.* **98**, 020403 (2007).
- [5] A. Bisio, G. Chiribella, G. M. D'Ariano, S. Facchini, and P. Perinotti, "Optimal quantum tomography of states, measurements, and transformations," *Phys. Rev. Lett.* **102**, 010404 (2009).
- [6] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y. K. Liu, "Efficient quantum state tomography," *Nat. Commun.* **1**, 149 (2010).
- [7] D. Gross, Y. K. Liu, S. T. Flammia, S. Becker, and J. Eisert, "Quantum state tomography via compressed sensing," *Phys. Rev. Lett.* **105**, 150401 (2010).
- [8] J. Řeháček, D. Mogilevtsev, and Z. Hradil, "Operational tomography: fitting of data patterns," *Phys. Rev. Lett.* **105**, 010402 (2010).
- [9] G. Tóth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, "Permutationally invariant quantum tomography," *Phys. Rev. Lett.* **105**, 250403 (2010).
- [10] W. T. Liu, T. Zhang, J. Y. Liu, P. X. Chen, and J. M. Yuan, "Experimental quantum state tomography via compressed sampling," *Phys. Rev. Lett.* **108**, 170403 (2012).
- [11] M. Christandl and R. Renner, "Reliable quantum state tomography," *Phys. Rev. Lett.* **109**, 120403 (2012).
- [12] J. S. Lundeen and C. Bamber, "Procedure for direct measurement of general quantum states using weak measurement," *Phys. Rev. Lett.* **108**, 070402 (2012).
- [13] J. Z. Salvail, M. Agnew, A. S. Johnson, E. Bolduc, J. Leach, and R. W. Boyd, "Characterization of polarization states of light via direct measurement," *Nat. Photonics* **7**, 316 (2013).
- [14] J. B. Altepeter, E. R. Jeffrey, and P. G. Kwiat, "Photonic state tomography," *Advances in Atomic, Molecular, and Optical Physics* **52**, 105159 (2005).
- [15] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, "Measurement of qubits," *Phys. Rev. A* **64**, 052312 (2001).
- [16] J. Řeháček, Z. Hradil, and M. Ježek, "Iterative algorithm for reconstruction of entangled states" *Phys. Rev. A* **63**, 040303(R) (2001).
- [17] R. Blume-Kohout, "Hedged maximum likelihood quantum state estimation," *Phys. Rev. Lett.* **105**, 200504 (2010).
- [18] Y. S. Teo, H. Zhu, B. G. Englert, J. Řeháček, and Z. Hradil, "Quantum-state reconstruction by maximizing likelihood and entropy," *Phys. Rev. Lett.* **107**, 020404 (2011).
- [19] Y. S. Teo, B. Stoklasa, B. G. Englert, J. Řeháček, and Z. Hradil, "Incomplete quantum state estimation: a comprehensive study," *Phys. Rev. A* **85**, 042317 (2012).
- [20] J. A. Smolin, J. M. Gambetta, and G. Smith, "Efficient method for computing the maximum-likelihood quantum state from measurements with additive Gaussian noise," *Phys. Rev. Lett.* **108**, 070502 (2012).
- [21] E. Halenková, K. Lemr, A. Černocho, and J. Soubusta, "Experimental simulation of a polarization-dispersion-fluctuating channel with photon pairs," *Phys. Rev. A* **85**, 063807 (2012).
- [22] E. Halenková, A. Černocho, K. Lemr, J. Soubusta, and S. Drusová, "Experimental implementation of the multifunctional compact two-photon state analyzer," *Appl. Opt.* **51**, 474 (2012).
- [23] J. Peřina Jr., M. Hamar, V. Michálek, and O. Haderka, "Photon-number distributions of twin beams generated in spontaneous parametric down-conversion and measured by an intensified CCD camera," *Phys. Rev. A* **85**, 023816 (2012).
- [24] K. Lemr, K. Bartkiewicz, A. Černocho, J. Soubusta, and A. Miranowicz, "Experimental linear-optical implementation of a multifunctional optimal cloner," *Phys. Rev. A* **85**, 050307(R) (2012).
- [25] K. Bartkiewicz, K. Lemr, A. Černocho, J. Soubusta, and A. Miranowicz, "Experimental eavesdropping based on optimal quantum cloning," *Phys. Rev. Lett.* **110**, 173601 (2013).
- [26] K. Bartkiewicz, A. Černocho, K. Lemr, J. Soubusta, and M. Stobińska, "Efficient amplification of photonic qubits by optimal quantum cloning" *Phys. Rev. A* **89**, 062322 (2014).
- [27] T. Opatrny, D.-G. Welsch, and W. Vogel, "Least-squares inversion for density-matrix reconstruction," *Phys. Rev. A* **56**, 1788 (1997).
- [28] J. Peřina Jr., O. Haderka, V. Michálek, and M. Hamar, "State reconstruction of a multimode twin beam using photodetection," *Phys. Rev. A* **87**, 022108 (2013).
- [29] R. Blume-Kohout, "Optimal reliable estimation of quantum states," *New J. Phys.* **12**, 043034 (2010).
- [30] F. Huszár and N. M. T. Houlby, "Adaptive Bayesian quantum tomography" *Phys. Rev. A* **85**, 052120 (2012).
- [31] B. Qi, Z. Hou, L. Li, D. Dong, G. Xiang, and G. Guo, "Quantum state tomography via linear regression estimation," *Sci. Rep.* **3**, 3496 (2013).
- [32] A. Roy and A. J. Scott, "Weighted complex projective 2-designs from bases: Optimal state determination by orthogonal measurements," *J. Math. Phys.* **48**, 072110 (2007).
- [33] A. B. Klimov, G. Björk, and L. L. Sánchez-Soto, "Optimal quantum tomography of permutationally invariant qubits," *Phys. Rev. A* **87**, 012109 (2013).
- [34] W. K. Wootters and B. D. Fields, "Optimal state-determination by mutually unbiased measurements," *Ann. Phys.* **191**, 363 (1989).
- [35] J. Řeháček, B.-G. Englert, and D. Kaszlikowski, "Minimal qubit tomography," *Phys. Rev. A* **70**, 052321 (2004).

- [36] G. M. D'Ariano, P. Mataloni, and M. F. Sacchi, "Generating qudits with  $d=3,4$  encoded on two-photon states," *Phys. Rev. A* **71**, 062337 (2005).
- [37] A. Ling, K. P. Soh, A. Lamas-Linares, and C. Kurtsiefer, "Experimental polarization state tomography using optimal polarimeters," *Phys. Rev. A* **74**, 022309 (2006).
- [38] M. D. de Burgh, N. K. Langford, A. C. Doherty, and A. Gilchrist, "Choice of measurement sets in qubit tomography," *Phys. Rev. A* **78**, 052122 (2008).
- [39] R. B. A. Adamson and A. M. Steinberg, "Improving quantum state estimation with mutually unbiased bases," *Phys. Rev. Lett.* **105**, 030406 (2010).
- [40] J. Nunn, B. J. Smith, G. Puentes, I. A. Walmsley, and J. S. Lundeen, "Optimal experiment design for quantum state tomography: fair, precise, and minimal tomography," *Phys. Rev. A* **81**, 042109 (2010).
- [41] S. N. Filippov and V. I. Man'ko, "Inverse spin-s portrait and representation of qudit states by single probability vectors," *J. Russian Laser Res.* **31**, 32 (2010).
- [42] Yu. I. Bogdanov, G. Brida, M. Genovese, S. P. Kulik, E. V. Moreva, and A. P. Shurupov, "Statistical estimation of the efficiency of quantum state tomography protocols," *Phys. Rev. Lett.* **105**, 010404 (2010).
- [43] Yu. I. Bogdanov, S. P. Kulik, E. V. Moreva, I. V. Tikhonov, and A. K. Gavrichenk, "Optimization of a Quantum Tomography Protocol for Polarization Qubits," *JETP Letters* **91**, 686 (2010).
- [44] Yu. I. Bogdanov, G. Brida, I. D. Bukeev, M. Genovese, K. S. Kravtsov, S. P. Kulik, E. V. Moreva, A. A. Soloviev, and A. P. Shurupov, "Statistical estimation of the quality of quantum-tomography protocols," *Phys. Rev. A* **84**, 042108 (2011).
- [45] K. E. Atkinson, *An Introduction to Numerical Analysis* (Wiley, New York, 1989).
- [46] N. J. Higham, *Accuracy and Stability of Numerical Algorithms* (SIAM, Philadelphia, 1996).
- [47] G. Golub and C. F. van Loan, *Matrix Computations* (Johns Hopkins University Press, Baltimore, 1989).
- [48] W. Kahan, "Numerical linear algebra," *Canad. Math. Bull.* **9**, 757 (1966).
- [49] G. Kimura, "The Bloch vector for  $N$ -level systems," *Phys. Lett. A* **314**, 339 (2003).
- [50] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury, and F. Vatan, "A new proof for the existence of mutually unbiased bases," *Algorithmica* **34**, 512 (2002), arXiv:quant-ph/0103162v3.
- [51] P. K. Aravind, "Solution to the king's problem in prime power dimensions," *Z. Naturforsch.* **58a**, 85 (2003).
- [52] J. B. Altepeter, E. R. Jeffrey, and P. G. Kwiat, "Phase-compensated ultra-bright source of entangled photons," *Opt. Express* **13**, 8951 (2005).
- [53] G. L. Long, H. Y. Yan, and Y. Sun, "Analysis of density matrix reconstruction in NMR quantum computing," *J. Opt. B: Quantum Semiclass. Opt.* **3**, 376 (2001).
- [54] A. M. Turing, "Rounding-off errors in matrix processes," *Quart. J. Mech. Appl. Math.* **1**, 287 (1948).
- [55] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra* (SIAM, Philadelphia, 2000).
- [56] X. Ji and B. H. Wildenthal, "Effective interaction for  $N = 50$  isotones," *Phys. Rev. C* **37**, 1256 (1988).
- [57] P. Kok and B. W. Lovett, *Introduction to Optical Quantum Information Processing* (Cambridge University Press, Cambridge, 2010).
- [58] M. Bartkowiak and A. Miranowicz, "Linear-optical implementations of the iSWAP and controlled NOT gates based on conventional detectors," *J. Opt. Soc. Am. B* **27**, 2369 (2010).
- [59] E. Knill, R. Laflamme, and G. J. Milburn, "A scheme for efficient quantum computation with linear optics," *Nature (London)* **409**, 46 (2001).
- [60] M. Koashi, T. Yamamoto, and N. Imoto, "Probabilistic manipulation of entangled photons," *Phys. Rev. A* **63**, 030301(R) (2001).
- [61] Y. Hirayama, A. Miranowicz, T. Ota, G. Yusa, K. Muraki, S. K. Özdemir, and N. Imoto, "Nanometre-scale nuclear-spin device for quantum information processing," *J. Phys.: Condens. Matter* **18**, S885 (2006).
- [62] K. Bartkiewicz, K. Lemr, A. Černoč, J. Soubusta, and A. Miranowicz, in preparation.
- [63] S. K. Özdemir, A. Miranowicz, M. Koashi, and N. Imoto, "Quantum scissors device for optical state truncation: A proposal for practical realization," *Phys. Rev. A* **64**, 063818 (2001).
- [64] J. Patera and H. Zassenhaus, "The Pauli matrices in  $n$  dimensions and finest gradings of simple Lie algebras of type  $A_{n-1}$ ," *J. Math. Phys.* **29**, 665 (1988).
- [65] P. Skrzypczyk, M. Navascués, and D. Cavalcanti, "Quantifying Einstein-Podolsky-Rosen steering," *Phys. Rev. Lett.* **112**, 180404 (2014).
- [66] K. Bartkiewicz, K. Lemr, A. Černoč, and J. Soubusta, "Measuring nonclassical correlations of two-photon states," *Phys. Rev. A* **87**, 062102 (2013).